

1648#13



1600

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/029,840A

DATE: 10/16/2002

TIME: 15:32:10

RECEIVED

OCT 25 2002

Input Set : A:\AM100389-US.ST25.txt

Output Set: N:\CRF4\10162002\J029840A.raw

TECH CENTER 1600/2900

3 <110> APPLICANT: Meng, Xiang-Jin
 4 Haqshenas, Gholamreza
 5 Huang, Fang-Fang
 7 <120> TITLE OF INVENTION: AVIAN HEPATITIS E VIRUS, VACCINES AND METHODS OF PROTECTING
 8 AGAINST AVIAN HEPATITIS-SPLENOMEGALY SYNDROME AND MAMMALIAN
 9 HEPATITIS E
 11 <130> FILE REFERENCE: AM100389
 13 <140> CURRENT APPLICATION NUMBER: 10/029840A
 14 <141> CURRENT FILING DATE: 2001-12-31
 16 <160> NUMBER OF SEQ ID NOS: 25
 18 <170> SOFTWARE: PatentIn version 3.1

ENTERED

20 <210> SEQ ID NO: 1
 21 <211> LENGTH: 3946
 22 <212> TYPE: DNA
 23 <213> ORGANISM: Hepatitis E virus
 25 <400> SEQUENCE: 1

26	accagcattg	gatttcgatg	gacgctgttt	aacgagcgcc	gttgatcttg	ggttgagcc	60
28	taccagctgg	cgcaccgtat	cccaccgttg	cccttgggac	gtttgtatat	ttttgcgtac	120
30	tgattatccg	actatcacca	caaccagtag	ggtgctgcgg	tctgttgtgt	ttaccgggtga	180
32	aaccattggt	cagaagatag	tgtttaccca	ggtggccaag	cagtcgaacc	ccgggtccat	240
34	aacgggtccat	gaggcgcagg	gcagtacttt	tgatcagact	actataatcg	ccacgttaga	300
36	tgctcgtggc	cttatagctt	catctcgcgc	gcatgccata	gttgcgctaa	cccgccaccg	360
38	ggagcgtgt	agtgtgattg	atgttggtgg	ggtgctggtc	gagattggag	ttactgatgc	420
40	catgtttaac	aatatcgaaa	tcagccttgt	gcgacctgat	gctgcagccc	ctgccgggggt	480
42	gctacgagcc	ccagacgaca	ccgtggatgg	cttgttgagc	atacccccg	cccacactga	540
44	tgtagcggcg	gtgttaacag	ctgaggcgat	tgggcatgcg	ccccttgaat	tggccgccat	600
46	aaatccaccc	gggcctgtat	tggagcaggg	cctattatac	atgccggcca	ggcttgatgg	660
48	gcgtgatgag	gttggttaagc	tccagctgtc	ggatactgta	cactgccgcc	tggctgcacc	720
50	cactagccgt	cttgccgtga	ttaacacatt	ggttgggcgg	tacggtaaag	ccactaagct	780
52	gcctgagggt	gaatatgact	taatggacac	tattgcgcag	ttctggcatc	atatcggacc	840
54	aatcaacccc	tcaacactgg	agtatgcaga	gatgtgcgag	gccatgctta	gtaaggcca	900
56	ggatgggtcc	ttgattgtac	atctggattt	acaggatgct	gattgttctc	gcataacatt	960
58	cttccagaag	gactgcgcta	aatttacgct	ggatgaccct	gttgcacacg	gtaaagtggg	1020
60	acaggggata	tctgcgtggc	cgaaaacttt	gtgtgcactt	ttcggcccct	ggttccgggc	1080
62	tatagagaag	cacctgtgtg	ctgggttacc	ccaggttat	tactatgggg	acctgtacac	1140
64	ggaagcogat	ctgcatcggt	ctgtgctttg	cgcgcctgct	ggtcaccttg	tttttgagaa	1200
66	tgattttcca	gagtttgact	caacgcagaa	taatgtgtcc	cttgatctcg	aatgtgaatt	1260
68	gatgcgcagg	tttgggatgc	ccgattggat	ggtagccttg	taccatcttg	ttcgatcata	1320
70	ctggctcttg	gttgccccga	agaagccct	tcgtggctgt	tggaaaaaac	actctggtga	1380
72	gcggggcacc	cttttgtgga	atacagtttg	gaacatgact	gtgttgcatc	atgtttatga	1440
74	gtttgatoga	ccaagtgtgt	tgtgtttcaa	aggtgatgat	agtgtcgttg	tctgtgaatc	1500
76	ggtgcgcgcc	cgtccagagg	gcgttagtct	cgtggcagac	tgccgggctaa	aaatgaagga	1560
78	caagaccggc	ccgtgtggcg	ccttttccaa	cctgctgatc	ttcccgggag	ctggtgttgt	1620

RAW SEQUENCE LISTING

DATE: 10/16/2002

PATENT APPLICATION: US/10/029,840A

TIME: 15:32:10

Input Set : A:\AM100389-US.ST25.txt

Output Set : N:\CRF4\10162002\J029840A.raw

```

80 ctgcgacctg ttacggcagt ggggcccgtt gactgacaag aactgggggc cgcacattca 1680
82 gcggatgcag gaccttgagc aagcgtgtaa ggattttgtt gcacgtgttg taactcaggg 1740
84 taaagagatg ttgaccatcc agcttgtggc gggttattat ggtgtggaag ttggtatggt 1800
86 tgaggtgggt tggggggcct tgaaggcctg cgcgcagacc cgcgagacc tagtgacca 1860
88 caggttgccg gtactaaact tatctaagga ggactgaaca aataacaatc attatgcagt 1920
90 ctgcgcgtcc atgtgcctta gctgccagtt ctggtgtttg gagtgccagg aaagtggggt 1980
92 gggatgtcgc tgtgtagatt gttgctcatg cttgcaatgt gctgcggggt gtcaaggggc 2040
94 tcccaaacgc tccagccgg aggaggcgt ggccagcgcc gccgtgacaa ttcagcccag 2100
96 tggagcactc aacaacgccc cgaggagacc gtcggccccg cccctctcac agacgttgct 2160
98 accgcgccag gtactcgcac ggtaccagat gtagatcaag ccggtgccgt gctggtgcgc 2220
100 cagtataatc tagtgaccag cccgttaggc ctggccaccc ttggtagcac caatgccttg 2280
102 ctttatgccg caccggtgtc accgttaatg ccgcttcagg acggcacgac gtctaatac 2340
104 atgagcacgg agtctagcaa ctatgctcaa taccgtgtac agggcctaac tgtccgctgg 2400
106 cgcgcagttg tgccaaatgc ggtgggcggc ttctctataa gcatggccta ttggccccag 2460
108 acaacatcca cccctacaag cattgacatg aattccatca cgtccactga cgtccgtgtg 2520
110 gtgcttcagc cgggctctgc tggtttgctg actataccac atgagcgttt ggcgtataag 2580
112 aacaatggtt ggcggtccgt cgaaacggtt tccgtccac aggaggatgc cagtcccggc 2640
114 atgctcatgg tttgtgtcca cgggaccccc tgggaatagt ataccaatag tgtttacacc 2700
116 gggccgcttg gtatggttga ttttgccata agtttacagc taaggaaact gtcgcccggg 2760
118 aatacaaatg ccagggtcac ccgtgtgaag gtgacggccc cacataccat caaggctgac 2820
120 ccattctggt ctaccataac aacagcagct gcggccaggt ttatggcgga tgtgcgttgg 2880
122 ggcttgggca ctgctgagga tggcgaaatt ggtcacggca tccttggtgt tctgtttaac 2940
124 ctggcgga caagtcttag tggcttgccc tcgacactgc tgcgggcggc gagtggtcag 3000
126 tacatgtacg gccggcctgt ggggaacgcg aacggcgagc ctgaggtgaa actgtatatg 3060
128 tcggttgagg atgcggttaa cgataaacct attatggtcc cccatgacat cgacctcggg 3120
130 accagcactg tcacctgcca ggactatggg aatcagcatg tggatgaccg cccatccccg 3180
132 gccccggccc ctaagcgagc tttgggcacc ctaaggtcag gggatgtgtt gcgtattact 3240
134 ggcctccatg agtatgtgac taacgcccag ttgttaccgc agagtgtgtc acaggggtac 3300
136 tttggggccg gcagcaccat gatggtgcat aatttgatca ctggtgtgcg cgcgcccgcc 3360
138 agttcagtcg actggacgaa ggcaacagtg gatgggtcc aggtgaagac tgtcgatgct 3420
140 agttctggga gtaataggtt tgcagcgtaa cctgcatttg gaaagccagc tgtgtggggg 3480
142 cccagggcg ctgggtattt ctaccagtat aacagcacc accaggagtg gatttatttt 3540
144 cttcagaatg gtagctccgt ggtttggtat gcatatacta atatgttggg ccagaagtca 3600
146 gatacatcca ttctttttga ggtccggcca atccaagcta gtgatcagcc ttggtttttg 3660
148 gcacaccaca ctggcgcgga tgactgtacc acctgtctgc ctctggggtt aagaacatgt 3720
150 tgccgccagg cgccagaaga ccagtcacct gagacgcgcc ggctcctaga ccggcttagt 3780
152 aggacattcc cctcaccacc ctaatgtcgt ggttttgggg ttttaggttg attttctgta 3840
154 tctgggcgta attgccccta tgtttaattt attgtgattt ttataactgt tcatttgatt 3900
156 atttatgaaa tcctcccatc tcgggcatag taaaaaaaa aaaaaa 3946

```

159 <210> SEQ ID NO: 2

160 <211> LENGTH: 146

161 <212> TYPE: PRT

162 <213> ORGANISM: Hepatitis E virus

164 <400> SEQUENCE: 2

```

166 Pro Ala Leu Asp Phe Asp Gly Arg Cys Leu Thr Ser Ala Val Asp Leu
167 1          5          10          15
170 Gly Leu Gln Pro Thr Ser Trp Arg Thr Val Ser His Arg Cys Pro Trp
171          20          25          30
174 Asp Val Cys Ile Phe Leu Arg Thr Asp Tyr Pro Thr Ile Thr Thr Thr

```

RAW SEQUENCE LISTING

DATE: 10/16/2002

PATENT APPLICATION: US/10/029,840A

TIME: 15:32:10

Input Set : A:\AM100389-US.ST25.txt

Output Set: N:\CRF4\10162002\J029840A.raw

```

175          35          40          45
178 Ser Arg Val Leu Arg Ser Val Val Phe Thr Gly Glu Thr Ile Gly Gln
179          50          55          60
182 Lys Ile Val Phe Thr Gln Val Ala Lys Gln Ser Asn Pro Gly Ser Ile
183 65          70          75          80
186 Thr Val His Glu Ala Gln Gly Ser Thr Phe Asp Gln Thr Thr Ile Ile
187          85          90          95
190 Ala Thr Leu Asp Ala Arg Gly Leu Ile Ala Ser Ser Arg Ala His Ala
191          100          105          110
194 Ile Val Ala Leu Thr Arg His Arg Glu Arg Cys Ser Val Ile Asp Val
195          115          120          125
198 Gly Gly Val Leu Val Glu Ile Gly Val Thr Asp Ala Met Phe Asn Asn
199          130          135          140
202 Ile Glu
203 145
206 <210> SEQ ID NO: 3
207 <211> LENGTH: 439
208 <212> TYPE: DNA
209 <213> ORGANISM: Hepatitis E virus
211 <400> SEQUENCE: 3
212 accagcattg gatttcgatg gacgctgttt aacgagcgcc gttgatcttg ggttcagcc 60
214 taccagctgg cgcaccgtat cccaccgttg cccttgggac gtttgtatat ttttcgtac 120
216 tgattatccg actatcacca caaccagtag ggtgctgcgg tctgttgtgt ttaccggtga 180
218 aaccattggt cagaagatag tgtttaccca ggtggccaag cagtcgaacc ccgggtccat 240
220 aacgggtccat gaggcgcagg gcagtacttt tgatcagact actataatcg ccacgttaga 300
222 tgctcgtggc cttatagctt catctcgcgc gcatgccata gttgcgctaa cccgccaccg 360
224 ggagcgtgt agtgtgattg atgttggtgg ggtgctggtc gagattggag ttactgatgc 420
226 catgtttaac aatatcgaa 439
229 <210> SEQ ID NO: 4
230 <211> LENGTH: 483
231 <212> TYPE: PRT
232 <213> ORGANISM: Hepatitis E virus
234 <400> SEQUENCE: 4
236 Leu Val Arg Pro Asp Ala Ala Ala Pro Ala Gly Val Leu Arg Ala Pro
237 1          5          10          15
240 Asp Asp Thr Val Asp Gly Leu Leu Asp Ile Pro Pro Ala His Thr Asp
241          20          25          30
244 Val Ala Ala Val Leu Thr Ala Glu Ala Ile Gly His Ala Pro Leu Glu
245          35          40          45
248 Leu Ala Ala Ile Asn Pro Pro Gly Pro Val Leu Glu Gln Gly Leu Leu
249          50          55          60
252 Tyr Met Pro Ala Arg Leu Asp Gly Arg Asp Glu Val Val Lys Leu Gln
253 65          70          75          80
256 Leu Ser Asp Thr Val His Cys Arg Leu Ala Ala Pro Thr Ser Arg Leu
257          85          90          95
260 Ala Val Ile Asn Thr Leu Val Gly Arg Tyr Gly Lys Ala Thr Lys Leu
261          100          105          110
264 Pro Glu Val Glu Tyr Asp Leu Met Asp Thr Ile Ala Gln Phe Trp His
265          115          120          125

```

RAW SEQUENCE LISTING

DATE: 10/16/2002

PATENT APPLICATION: US/10/029,840A

TIME: 15:32:10

Input Set : A:\AM100389-US.ST25.txt

Output Set: N:\CRF4\10162002\J029840A.raw

```

268 His Ile Gly Pro Ile Asn Pro Ser Thr Leu Glu Tyr Ala Glu Met Cys
269      130      135      140
272 Glu Ala Met Leu Ser Lys Gly Gln Asp Gly Ser Leu Ile Val His Leu
273 145      150      155      160
276 Asp Leu Gln Asp Ala Asp Cys Ser Arg Ile Thr Phe Phe Gln Lys Asp
277      165      170      175
280 Cys Ala Lys Phe Thr Leu Asp Asp Pro Val Ala His Gly Lys Val Gly
281      180      185      190
284 Gln Gly Ile Ser Ala Trp Pro Lys Thr Leu Cys Ala Leu Phe Gly Pro
285      195      200      205
288 Trp Phe Arg Ala Ile Glu Lys His Leu Val Ala Gly Leu Pro Pro Gly
289      210      215      220
292 Tyr Tyr Tyr Gly Asp Leu Tyr Thr Glu Ala Asp Leu His Arg Ser Val
293 225      230      235      240
296 Leu Cys Ala Pro Ala Gly His Leu Val Phe Glu Asn Asp Phe Ser Glu
297      245      250      255
300 Phe Asp Ser Thr Gln Asn Asn Val Ser Leu Asp Leu Glu Cys Glu Leu
301      260      265      270
304 Met Arg Arg Phe Gly Met Pro Asp Trp Met Val Ala Leu Tyr His Leu
305      275      280      285
308 Val Arg Ser Tyr Trp Leu Leu Val Ala Pro Lys Glu Ala Leu Arg Gly
309      290      295      300
312 Cys Trp Lys Lys His Ser Gly Glu Pro Gly Thr Leu Leu Trp Asn Thr
313 305      310      315      320
316 Val Trp Asn Met Thr Val Leu His His Val Tyr Glu Phe Asp Arg Pro
317      325      330      335
320 Ser Val Leu Cys Phe Lys Gly Asp Asp Ser Val Val Val Cys Glu Ser
321      340      345      350
324 Val Arg Ala Arg Pro Glu Gly Val Ser Leu Val Ala Asp Cys Gly Leu
325      355      360      365
328 Lys Met Lys Asp Lys Thr Gly Pro Cys Gly Ala Phe Ser Asn Leu Leu
329      370      375      380
332 Ile Phe Pro Gly Ala Gly Val Val Cys Asp Leu Leu Arg Gln Trp Gly
333 385      390      395      400
336 Arg Leu Thr Asp Lys Asn Trp Gly Pro Asp Ile Gln Arg Met Gln Asp
337      405      410      415
340 Leu Glu Gln Ala Cys Lys Asp Phe Val Ala Arg Val Val Thr Gln Gly
341      420      425      430
344 Lys Glu Met Leu Thr Ile Gln Leu Val Ala Gly Tyr Tyr Gly Val Glu
345      435      440      445
348 Val Gly Met Val Glu Val Val Trp Gly Ala Leu Lys Ala Cys Ala Ala
349      450      455      460
352 Ala Arg Glu Thr Leu Val Thr Asn Arg Leu Pro Val Leu Asn Leu Ser
353 465      470      475      480
356 Lys Glu Asp
360 <210> SEQ ID NO: 5
361 <211> LENGTH: 1450
362 <212> TYPE: DNA
363 <213> ORGANISM: Hepatitis E virus

```

RAW SEQUENCE LISTING

DATE: 10/16/2002

PATENT APPLICATION: US/10/029,840A

TIME: 15:32:10

Input Set : A:\AM100389-US.ST25.txt

Output Set: N:\CRF4\10162002\J029840A.raw

365 <400> SEQUENCE: 5

```

366 gcttgtgcga cctgatgctg cagccccctgc cgggggtgcta cgagccccag acgacaccgt      60
368 ggatggcttg ttggacatac cccccggcca cactgatgta gcggcggtgt taacagctga      120
370 ggcgattggg catgcgcccc ttgaattggc cgccataaat ccaccggggc ctgtattgga      180
372 gcagggccta ttatacatgc cggccaggct tgatgggcgt gatgaggttg ttaagctcca      240
374 gctgtcggat actgtacact gccgcctggc tgcaccact agccgtcttg cggtgattaa      300
376 cacattgggt gggcggtacg gtaaagccac taagctgcct gaggttgaat atgacttaat      360
378 ggacactatt gcgcagttct ggcacatata cggaccaatc aaccctcaa cactggagta      420
380 tgcagagatt tgcgaggcca tgcttagtaa gggccaggat gggtccttga ttgtâcatct      480
382 ggatttacag gatgctgatt gttctcgcac aacattcttc cagaaggact gcgctaaatt      540
384 tacgctggat gacctgttg cacacggtaa agtgggacag gggatatctg cgtggccgaa      600
386 aacttttgtt gcacttttcg gcccctggtt ccgggctata gagaagcacc ttgtggctgg      660
388 gttacccccca ggttattact atggggacct gtacacggaa gccgatctgc atcgttctgt      720
390 gctttgcgcg cctgctggtc acctgtttt tgagaatgat ttctcagagt ttgactcaac      780
392 gcagaataat gtgtcccttg atctcgaatg tgaattgatg cgcaggtttg ggatgcccgga      840
394 ttggatggta gccttgatcc atctgttcg atcatactgg ctcttggttg ccccgaaaga      900
396 agcccttcgt ggctgttgga aaaaacactc tggtgagccg ggcacccttt tgtggaatac      960
398 agtttggaac atgactgtgt tgcacatgtt ttatgagttt gatcgaccaa gtgtgttggt      1020
400 tttcaaaggt gatgatagtg tcgttgtctg tgaatcgggt cgcgcccgtc cagagggcgt      1080
402 tagtctcgtg gcagactgcg ggctaaaaat gaaggacaag accggcccgt gtggcgccct      1140
404 ttccaacctg ctgatcttcc cgggagctgg tgttgtctgc gacctgttac ggcagtgggg      1200
406 ccgcttgact gacaagaact gggggcccga cattcagcgg atgcaggacc ttgagcaagc      1260
408 gtgtaaggat tttgttgac gtgttgtaac tcagggtaaa gagatgttga ccatccagct      1320
410 tgtggcgggt tattatggtg tggaagttgg tatggttgag gtggtttggg gggctttgaa      1380
412 ggctgcgcc gcagcccgcg agaccctagt gaccaacagg ttgccggtac taaacttacc      1440
414 taaggaggac

```

417 <210> SEQ ID NO: 6

418 <211> LENGTH: 606

419 <212> TYPE: PRT

420 <213> ORGANISM: Hepatitis E virus

422 <400> SEQUENCE: 6

```

424 Met Ser Leu Cys Arg Leu Leu Leu Met Leu Ala Met Cys Cys Gly Val
425 1          5          10          15
428 Ser Arg Gly Ser Gln Thr Leu Pro Ala Gly Gly Arg Arg Gly Gln Arg
429          20          25          30
432 Arg Arg Asp Asn Ser Ala Gln Trp Ser Thr Gln Gln Arg Pro Glu Gly
433          35          40          45
436 Ala Val Gly Pro Ala Pro Leu Thr Asp Val Val Thr Ala Ala Gly Thr
437          50          55          60
440 Arg Thr Val Pro Asp Val Asp Gln Ala Gly Ala Val Leu Val Arg Gln
441 65          70          75          80
444 Tyr Asn Leu Val Thr Ser Pro Leu Gly Leu Ala Thr Leu Gly Ser Thr
445          85          90          95
448 Asn Ala Leu Leu Tyr Ala Ala Pro Val Ser Pro Leu Met Pro Leu Gln
449          100         105         110
452 Asp Gly Thr Thr Ser Asn Ile Met Ser Thr Glu Ser Ser Asn Tyr Ala
453          115         120         125
456 Gln Tyr Arg Val Gln Gly Leu Thr Val Arg Trp Arg Pro Val Val Pro
457          130         135         140

```

VERIFICATION SUMMARY

DATE: 10/16/2002

PATENT APPLICATION: US/10/029,840A

TIME: 15:32:11

Input Set : A:\AM100389-US.ST25.txt

Output Set: N:\CRF4\10162002\J029840A.raw